Marta Małecka[*]

COMPARATIVE ANALYSIS OF SIGMA-BASED, QUANTILE-BASED AND TIME SERIES VAR ESTIMATORS

Abstract. Since its inception at the end of the XX century, VaR risk measure has gained massive popularity. It is synthetic, easy in interpretation and offers comparability of risk levels reported by different institutions. However, the crucial idea of comparability of reported VaR levels stays in contradiction with the differences in estimation procedures adopted by companies. The issue of the estimation method is subject to the internal company decision and is not regulated by the international banking supervision.

The paper was dedicated to comparative analysis of the prediction errors connected with competing VaR estimation methods. Four methods, among which two stationarity-based – variance-covariance and historical simulation – and two time series methods – GARCH and RiskMetrics[TM] – were compared through the Monte Carlo study. The analysis was conducted with respect to the method choice, series length and VaR tolerance level.

The study outcomes showed the superiority of the sigma-based method of variance-covariance over the quantile-based historical simulation. Furthermore the comparison of the stationarity-based estimates to the time series results showed that allowing for time-varying parameters in the estimation technique significantly reduces the estimator bias and variance.

Key words: VaR, VaR estimate, bias of the VaR estimator, variance of the VaR estimator, Monte Carlo experiment

1. INTRODUCTION

Being one of the most popular risk measures, VaR (value at risk) at the same time comes in for fierce criticism, especially after the financial crisis of the break of 2008 and 2009, which revealed serious shortages in existing risk management systems. Massive popularity of the VaR measure results from its clear interpretation and synthetic information about potential loss, contained in one number. From the point of view of the banking supervision, a key aspect is the possibility to apply VaR measure to various types of assets, portfolios and to compute the risk volume incurred by the whole institution, with proper regard to the diversification effect. As a consequence, it offers comparability of risk generated by institutions operating in the market, even if they run their activities in different market segments.

---

[*] Ph.D., Department of Statistical Methods, University of Łódź.

The crucial idea of comparability of reported VaR levels stays in contradiction with the differences in risk calculation procedures adopted by different institutions. Computation of VaR measure involves arbitrary elements, including tolerance level and estimation method. While the issue of the tolerance level has been regulated by the international banking supervision (Bank for International Settlements 1996, 2005a, 2005b), there are no specific recommendations about the estimation method.

Further problem connected with VaR estimation and comparability is the choice of the method of evaluating VaR forecasts. As VaR realizations are not observable, they cannot be compared to the forecasted values, thus the sample mean squared prediction error is not feasible. Analytical formulas for estimator variance cannot be applied in the situation when the assumption that the underlying variable distributions are identical over time is not fulfilled, which is common in the financial market.

VaR forecasts are commonly assessed with the use of statistical tests, however there is a broad discussion in literature about VaR tests power (cf. Berkowitz, Christoffersen, Pelletier 2011, Lopez 1999, Małecka 2013, Piontek 2014). In this paper we used simulation methods to compare the prediction errors connected with VaR estimation methods.

The aim of the paper was to evaluate and compare the bias and the variance of competing VaR estimators through the Monte Carlo study. We compared the statistical properties of four different estimation methods: based on stationarity assumption – variance covariance method and historical simulation – and two times series methods – GARCH and RiskMetrics[TM]. The comparative analysis referred also to different lengths of estimation window. The study included 1% and 5% VaR tolerance levels.

The paper consists of four sections. In the second section we define VaR and give the overview of VaR estimation methods. The third section presents the description of the simulation study and shows empirical results. The final section provides a summary and conclusion.

2. DEFINITION OF VAR AND VAR ESTIMATION METHODS

Let $(\Omega, \mathrm{F}, \mathrm{P})$ be the probabilistic space, where Ω is the set of all possible risk factors, F is the borel σ-algebra of Ω and P is the probability. Let F_t be the filtration generated by the subsets of Ω at time t. Let us denote $v : \mathbf{R}^k \mapsto \mathbf{R}$ the payoff function defined on the space of risk factor values, V_t – the

value of the payoff at time t: $V_t = v(P_t)$, where $P_t = [P_{t1}, P_{t2}, ..., P_{tk}]'$ is the vector of risk factors. Let X_t represent the profit and loss (P&L):

$$X_t = v(P_{t+1}) - v(P_t),\tag{1}$$

where $P_{t+1} = P_t \times \exp(R_t)$, $R_t = \left[\ln\dfrac{P_{t+1,1}}{P_{t1}}, \ln\dfrac{P_{t+1,2}}{P_{t2}}, ..., \ln\dfrac{P_{t+1,k}}{P_{tk}} \right]'$. Using the P&L variable, VaR at the tolerance level α is defined as its α-quantile:

$$VaR_{t,\alpha} = F_t^{-1}(\alpha),\tag{2}$$

where F_t is the distribution function of the P&L variable (cf. Bałamut 2002). Transformation of the above formula gives VaR in terms of the logarithmic rate of return R_t:

$$VaR_{t,\alpha} = V_t \left(\exp(\mu_t) - \exp(R_t^*) \right),\tag{3}$$

where R_t^* is the α-quantile of R_t and μ_t – its expectation. In a descriptive way VaR may be defined as the worst loss over a target horizon, such that there is a low, prespecified probability that the actual loss will be larger (Jorion 2007: 106–107).

VaR estimation methods can be divided into two basic groups: sigma-based methods and quantile-based methods (Jorion 1996). Under the assumption of the stationarity of the return series, the main representatives of these groups are methods called variance-covariance method and historical simulation, respectively. The sigma-based methods are based on the assumption that the α-quantile of the R_t distribution may be obtained as the function of the standard deviation. If we assume the model $R_t = \mu_t + z_t \sigma_t$, where z_t follows a standard distribution, we can get the quantile R_t^* as

$$R_t^* = \mu_t + z_\alpha \sigma_t \tag{4}$$

where z_α is the α-quantile of z_t. Thus R_t^* is the function of σ_t. In this approach therefore risk factors are summarized in the historical variance-covariance matrix. Substituting (4) into the VaR definition (3), we get

$$VaR_{t,\alpha} = V_t \exp(\mu_t)\left(1 - \exp(z_\alpha \sigma_t)\right) \qquad (5)$$

(cf. Jajuga 2000).

The main representative of quantile-based methods is the historical simulation method, where VaR estimate is based on the sample quantile. Historical data is used to compute the empirical distribution, through which VaR is computed according to the formula (3) (Grabowska 2000).

The numerical problem connected with historical simulation method is related to the choice of the series length. In case of long time series the approximate stationarity assumption is usually violated. Moreover, long estimation windows are not recommended in order to ensure relevant impact of new market information on risk estimates. On the other hand, short samples give very few observations in a 1% or 5% tail of the empirical distribution, which are typical tolerance levels in VaR analysis. This may result in discontinuous behaviour of estimators and large estimator variance. Thus, the observation number is a fundamental decision to be taken when employing historical simulation method to VaR estimation.

In contrast to sigma-based, quantile-based methods do not require any assumptions about the probability distributions and the stochastic structure of the underlying processes, which is regarded to be their key practical advantage. Furthermore, the quantile-based methods are easily implemented even in case of portfolios with complex financial instruments.

If we relax the stationarity assumption, VaR can be estimated through the wide class of time series methods. Commonly used in VaR empirical analysis are GARCH models and RiskMetrics[TM]. method. In the GARCH model it is assumed that logarithmic returns can be represented by the following mean equation

$$r_t = E(r_t \mid \Omega_{t-1}) + \varepsilon_t, \qquad (6)$$

where Ω_t is the set of all information at time t, $E(\varepsilon_t) = 0$, $\varepsilon_t = z_t \cdot \sqrt{h_t}$, $z_t \in N(0,1)$, $t = 1, 2, \ldots$, h_t is thus the conditional variance of ε_t. In the GARCH(p,q) process the conditional variance is a function of its past values and the past realizations of the error term (Bollerslev 1986):

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}. \qquad (7)$$

Let $\hat{r}_t$ and $\hat{h}_t$ denote the forecasts of the expected value and the variance of the logarithmic returns respectively. Then the quantile R_t^* is given by $\hat{r}_t + z_\alpha \sqrt{\hat{h}_t}$ and VaR is computed according to the following formula

$$VaR_\alpha = V_t \left(\exp(\hat{r}_t) - \exp(\hat{r}_t + z_\alpha \sqrt{\hat{h}_t}) \right). \qquad (8)$$

The GARCH model offers representation of the variance fluctuations over time and volatility clustering phenomenon. GARCH parameters are estimated by the quasi maximum likelihood method, which gives consistent estimators (Domański, Pruska 2000: 125–131).

In RiskMetrics[TM] (RM[TM]) the variance h_t is computed as the exponentially weighted moving average, hence the variance equation takes the form

$$h_t = (1-\lambda)\sum_{i=0}^{\infty} \lambda^i \left(r_{t-i-1} - \mu \right)^2, \qquad (9)$$

although in practice often the recursive formula is used:

$$h_t = (1-\lambda)(r_{t-1} - \mu)^2 + \lambda h_{t-1}, \qquad (10)$$

$z_t \in N(0,1)$, $t = 1, 2, \ldots$ For daily data it is often assumed that $\lambda = 0,94$, which means that the effective length of the estimation window is around 30 observations (Fiszeder 2009: 159). RM[TM]-based VaR is computed according to the same formula (8) as in case of the GARCH model.

In comparison with the GARCH model, RM[TM] is easier in implementation and its superiority over historical simulation method results from the fact that it assigns higher weights to the latest observations. The main criticism connected with RM[TM] is related to the arbitrary choice of the λ parameter. It is argued in literature that, if this method may be seen as the special case of the GARCH model, the λ parameter should be estimated (Fiszeder 2009: 159).

3. MONTE CARLO STUDY

The four VaR estimation methods: stationarity based variance-covariance and historical simulation and time series-based GARCH and RiskMetrics[TM] were compared through the Monte Carlo study. The study included the bias and

the variance of the considered VaR estimators. To represent the volatility clustering phenomenon and include time-varying parameters of the return distribution, the simulation experiment was based on the GARCH process. The number of replications was set to 10 000 and the experiment parameter values

$$r_t = \varepsilon_t \, , \ \varepsilon_t \in N(0, h_t)$$

$$\varepsilon_t = z_t \cdot \sqrt{h_t} \, , \ z_t \in N(0,1)$$

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}$$

$$10\,000 \left\{ \begin{array}{l} z_1 \\ z_2 \\ z_3 \\ \dots \\ z_{1000} \end{array} \right.$$

$$\left. \begin{array}{l} h_1 \\ h_2 = \omega + \alpha \varepsilon_1^2 + \beta h_1 \\ h_3 = \omega + \alpha \varepsilon_2^2 + \beta h_2 \\ \dots \\ h_{1000} = \omega + \alpha \varepsilon_{999}^2 + \beta h_{999} \end{array} \right\} 10\,000$$

$$\varepsilon_1 = z_1 \sqrt{h_1}$$
$$\varepsilon_2 = z_2 \sqrt{h_2}$$
$$\varepsilon_3 = z_3 \sqrt{h_3}$$
$$\dots$$
$$\varepsilon_{1000} = z_{1000} \sqrt{h_{1000}}$$

$$h_{1001} = \omega + \alpha \varepsilon_{1000}^2 + \beta h_{1000}$$

$$VaR_\alpha = 1 - \exp(z_\alpha \sqrt{h_{1001}})$$

Fig.1. Simulation experiment scheme

Source: own work.

were fixed through the initial study on the historical WIG20 data, based on the rolling regression with 1000-observation-long estimation window. From all obtained parameter sets two extreme cases were chosen – with the smallest and the highest value of the β parameter. The presample variance of the GARCH process was generated from the uniform distribution, whose minimum and maximum value were based on the historical WIG20 data. In variance sampling the number of replications was 10 000, which, together with GARCH simulation number gave 100 000 000 replications (Fig. 1).

The results of the study showed a large impact of the quantile order (or VaR tolerance level) on the evaluation of methods (Tab. 1, 2). For 5% VaR the bias results were not absolutely conclusive in relative evaluation of the variance-covariance and historical simulation methods – both results were convergent to a stable level of around 0.03 (Fig. 2). However, in case of 1% VaR, there was

a clear superiority of the variance-covariance method over the historical simulation (Fig. 3).

The variance of the estimators was lower for the variance-covariance method in case of both tolerance levels, with the differences for 1% results being more striking (Fig. 4, 5). The largest prevalence of the variance-covariance method in terms of the estimator variance was observed for shortest series (100 and 250 observations).

Table 1. Empirical bias and variance of 5% VaR estimators, experiment parameters:
$\omega = 0.000001$, $\alpha = 0.034873$, $\beta = 0.959960$

Estimation method	Estimation error	$T = 100$	$T = 250$	$T = 500$	$T = 750$	$T = 1000$
Variance-covariance	Bias	−0.0115	0.0041	0.0192	0.0279	0.0338
	Variance	0.0465	0.1308	0.2132	0.2499	0.2687
Historical simulation	Bias	0.0500	0.0292	0.0285	0.0306	0.0320
	Variance	0.1123	0.1605	0.2301	0.2623	0.2780

Source: own work.

Table 2. Empirical bias and variance of 1% VaR estimators, experiment parameters:
$\omega = 0.000001$, $\alpha = 0.034873$, $\beta = 0.959960$

Estimation method	Estimation error	$T = 100$	$T = 250$	$T = 500$	$T = 750$	$T = 1000$
Variance-covariance	Bias	-0.0163	0.0057	0.0272	0.0394	0.0478
	Variance	0.0929	0.2617	0.4264	0.5000	0.5374
Historical simulation	Bias	0.2014	0.1337	0.1520	0.1759	0.1940
	Variance	0.4228	0.4065	0.5322	0.6071	0.6492

Source: own work.

The comparison of the analysed methods in context of the series length showed a systematic increase in the bias of the variance-covariance estimator with lengthening the series from 250 to 1000 observations. The sharpest growth was between 250 and 500 observations, then the differences were milder. The bias results for the historical simulation did not exhibit such a regular behaviour. However, similarly to variance-covariance, the bias observed for the longest series oversized the bias for 250 observations and after that sample size the results were relatively stable. In case of both methods large bias (in absolute terms) was observed for the shortest considered series of 100 observations.
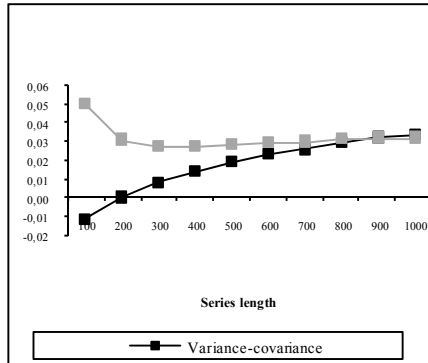
Fig. 2. Comparison of bias of variance-
covariance and historical simulation 5% VaR
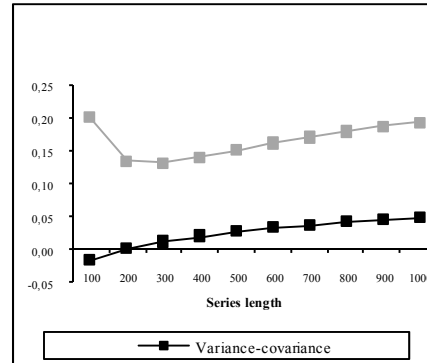estimators
Source: own work.

Fig. 3. Comparison of bias of variance-
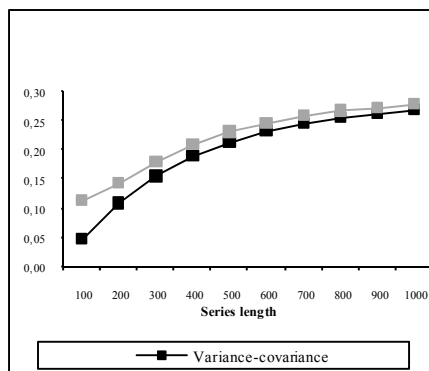covariance and historical simulation 1% VaR
estimators
Source: own work.

Fig. 4. Comparison of variance of variance-
covariance and historical simulation 5% VaR
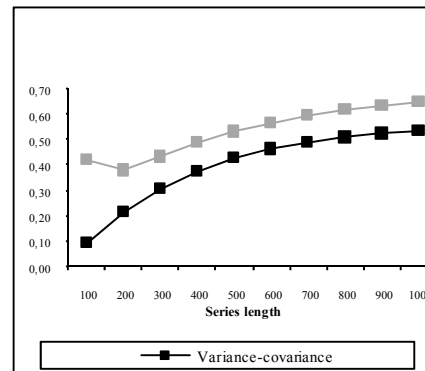estimators
Source: own work.

Fig. 5. Comparison of variance of variance-
covariance and historical simulation 1% VaR
estimators
Source: own work.

The variance results for different series lengths exhibited a more regular behaviour than bias estimates. Independent of the tolerance level and estimation method, the smallest variance was generally observed for the shortest series. Only for the historical simulation in the 1% VaR level estimation there was a slight drop with shift from 100 to 250 observations. In all other cases there was a clear rising tendency, with the highest increases observed for the shortest series. Over 500 observations the observed variance level was relatively stable.

The presented study confirmed the observation that the recommended series length for basic VaR estimation methods is 200–250 observations (Best 2000). Worsening statistical properties of the VaR estimators based on larger number of

observations, which seem contradictory to classic rules of statistical inference, are relevant for the processes with changing parameters. For long estimation window the assumption of the approximate stationarity is violated. Thus extending back the series brings no profit for the estimator quality.

On the other hand , shortening the series below 200 observations may also result in worse estimator properties, especially in case of the quantile-based method. For this method 1% VaR based on 100-observation series size is based merely on 1 observation and 5% VaR on 5 observations, which cannot ensure reliable estimates.

The comparison of the results observed for 1% and 5% VaR showed that there was a dependence of the obtained estimator properties on the chosen tolerance level. Independent of the estimation method, both in terms of the bias and the variance, the VaR estimates were more effective for the 5% tolerance level. This observation was more striking in case of the quantile-based method of historical simulation.

In the final part of the study the above results for the stationarity--assumption-based methods of variance-covariance and historical simulation were compared to the results obtained for the time series methods: GARCH and RiskMetrics^{TM}. Due to the sample size required for the time series methods, especially GARCH model estimated by the quasi maximum likelihood, the final comparison was restrained only to the series length of 1000 observations. The results of the study showed that relaxing the stationarity assumption and inclusion of the time-varying parameters allowed for a significant reduction of the estimator bias (Tab. 3, 4). Considering the small bias obtained for both GARCH and RM^{TM} estimates, the core part of the comparative analysis of the two methods was variance comparison. The GARCH-based estimator variance was over ten times lower than for RM^{TM} and was also much smaller than for both stationarity-based methods.

The simulation study presented in the paper was conducted for two variants of the simulation experiment, with two sets of the GARCH parameters. The results shown above were obtained for the parameters: $\omega = 0.000001$, $\alpha = 0.034873$, $\beta = 0.959960$. Performing the study for the second set of parameters: $\omega = 0.000027$, $\alpha = 0.168115$, $\beta = 0.781104$

did not alter the conclusions related to comparison of the methods or bias and variance analysis for different series lengths. The change in parameter values did not also influence the evaluation of methods in context of the tolerance level. The main difference was that the second set of parameters resulted in larger observed bias and variance, which may suggest that larger volatility persistence may improve the effectiveness of VaR estimates[1].

---

[1] For the sake of brevity the specific results about the bias and variance of the VaR estimators obtained for parameters $\omega = 0.000027$, $\alpha = 0.168115$, $\beta = 0.781104$ were not presented in the paper, however are available upon request from the author.

Table 3. Empirical bias and variance of 5% VaR estimators, experiment parameters:
$\omega = 0.000001$, $\alpha = 0.034873$, $\beta = 0.959960$

Estimation method	$T = 1000$	
	Bias	Variance
Variance-covariance	0.0345	0.2681
Historical simulation	0.0326	0.2773
GARCH	0.0003	0.0268
RM^{TM}	−0.0095	0.4420

Source: Own work.

Tab. 4. Empirical bias and variance of 5% VaR estimators, experiment parameters:
$\omega = 0.000001$, $\alpha = 0.034873$, $\beta = 0.959960$

Estimation method	$T = 1000$	
	Bias	Variance
Variance-covariance	0.0488	0.5362
Historical simulation	0.1951	0.6483
GARCH	0.0004	0.0536
RM^{TM}	−0.0135	0.8841

Source: Own work.

4. CONCLUSION

The presented study was dedicated to comparative analysis of the bias and the variance of VaR estimators. Four methods, among which two stationarity-based – variance-covariance and historical simulation – and two time series methods – GARCH and RiskMetrics^{TM} were compared through the Monte Carlo study. The analysis was conducted with respect to the method choice, series length and VaR tolerance level.

The study outcomes showed the superiority of the sigma-based method of variance-covariance over the quantile-based historical simulation, confirmed by the analysis of both the bias and the variance. The variance-covariance prevalence was larger for 1% VaR than for 5% tolerance level.

The presented results confirmed the observation that the series length of 200–250 observations offer better statistical properties than both shorter and longer series. This conclusion, contradictory to classic statistical inference, stays in line with the estimation procedure based on the stochastic process with time-varying parameters, where the assumption of the approximate stationarity is violated for the longer series. In case of changing process parameters extending the series does not bring improvement to estimation effectiveness.

Comparison of the stationarity-based estimates to the time series results showed that allowing for time-varying parameters in the estimation technique significantly reduces the estimator bias. In terms of the variance comparison, the GARCH-based VaR estimates outperformed both the RM^{TM} and stationarity-based methods.

REFERENCES

Basel Committee on Banking Supervision (1996), *Amendment to the capital accord to incorporate market risks*, online: http://www.bis.org/publ/bcbs24.pdf (accessed 20.12.2013).

Basel Committee on Banking Supervision (2005), *Amendment to the capital accord to incorporate market risks*, online: www.bis.org/publ/bcbs119.pdf (accessed 20.12.2013).

Basel Committee on Banking Supervision (2005), *Basel II: International Convergence of Capital Measurement and Capital Standards: a Revised Framework*, online: http://www.bis.org/publ/bcbs118.htm (accessed 20.12.2013).

Bałamut T. (2002), *Metody estymacji Value at Risk*, „Materiały i studia NBP" 147, 1-107.

Berkowitz J., Christoffersen P., Pelletier D. (2011), *Evaluating Value-at-Risk Models with Desk-Level Data*, Management Science" 12(57), 2213–2227.

Best P. (2000), *Wartość narażona na ryzyko*, Dom wydawniczy ABC, Kraków.

Bollerslev T. (1986), *Generalised Autoregressive Heteroscedasticity*, „Journal of Econometrics" 31, 307–327.

Domański Cz., Pruska K. (2000), *Nieklasyczne metody statystyczne*, PWE, Warszawa.

Fiszeder P. (2009), *Modele klasy GARCH w empirycznych badaniach finansowych*, Wydawnictwo Naukowe Uuniwersytetu Mikołaja Kopernika, Toruń.

Grabowska A. (2000), *Metody kalkulacji wartości narażonej na ryzyko (VaR)*, „Bank i kredyt" 32(10), 29–36.

Jajuga K. (2000), *Miary ryzyka rynkowego – część trzecia*, „Rynek Terminowy" 8, 112–117.

Jorion P. (1996), *Risk2: Measuring the Risk in Value at Risk*, „Financial Analysis Journal" 52(6), 47–56.

Jorion P. (2007), *Value at Risk. The New Benchmark for Managing Financial Risk*, McGraw-Hill.

Lopez J. (1999), *Methods for Evaluating Value-at-Risk Estimates*, „FRBSF Economic Review" 2, 3–17.

Małecka M. (2013), *Metody oceny jakości prognoz ryzyka rynkowego – analiza porównawcza*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu" 323, 192–201.

Piontek K. (2014), *Power analysis of some chosen tests of independence of Value-at-Risk violations*, International Conference Financial Investments and Insurance, September 17–19, 2014, Wrocław, Poland, online: http://www.inwest.ue.wroc.pl/index.php/en/programme (accessed 25.11.2014).

Marta Małecka

ANALIZA PORÓWNAWCZA ESTYMATORÓW VAR OPARTYCH NA WARIANCJI, NA METODACH KWANTYLOWYCH I METODACH SZEREGÓW CZASOWYCH

Streszczenie. Od czasu wprowadzenia VaR pod koniec XX wieku, miara ta stała się najpopularniejszą miarą ryzyka. Jako główne jej zalety uznaje się: łatwość interpretacji, możliwość uzyskania syntetycznej informacji o poziomie ryzyka w postaci jednaj liczby oraz porównywalność poziomów ryzyka raportowanych przez różne instytucje. Jednak możliwość porównywania poziomów ryzyka pozostaje w sprzeczności z faktem stosowania różnych procedur wyznaczania tej miary. Wybór metody estymacji jest wewnętrzną decyzją przedsiębiorstwa i nie podlega regulacjom międzynarodowego nadzoru bankowego.

Praca poświęcona została analizie porównawczej błędów estymatora związanych z konkurent-cyjnymi metodami szacowania VaR. Za pomocą badania Monte Carlo porównano cztery metody, wśród których wybrano dwie oparte na założeniu stacjonarności rozkładu – metodę wariancji-kowariancji oraz symulacji historycznej – oraz dwie metody szeregów czasowych – GARCH i RiskMetrics[TM]. Analiza porównawcza została przeprowadzona ze względu na wybór metody estymacji, długość szeregu czasowego oraz poziom tolerancji VaR.

Wyniki badania pokazały przewagę estymatorów VaR opartych na wariancji nad kwantylową metodą symulacji historycznej. Ponadto porównanie estymatorów opartych na założeniu stacjonarności z estymatorami wywodzącymi się z metod szeregów czasowych pokazało, że uwzględnienie zmienności parametrów pozwoliło na znaczącą redukcję obciążenia i wariancji estymatorów.

Słowa kluczowe: VaR, oszacowanie VaR, obciążenie estymatora VaR, wariancja estymatora VaR, eksperyment Monte Carlo.